Topological study and Lyapunov exponent of a secure steganographic scheme

Jacques Bahi, Nicolas Friot, Christophe Guyeux

Topological study and Lyapunov exponent of a secure steganographic scheme

Jacques M. Bahi[1], Nicolas Friot[1], and Christophe Guyeux[1*]
[1] *Computer science laboratory DISC*
FEMTO-ST Institute, UMR 6174 CNRS
University of Franche-Comté, Belfort, France
{*jacques.bahi, nicolas.friot, christophe.guyeux*}*@femto-st.fr*
[*] *Authors in alphabetic order*

Keywords: *Information Hiding, Steganography, Security, Topology, Lyapunov Exponent*

Abstract:
CIS_2 is a steganographic scheme proposed formerly, belonging into the small category of algorithms being both stego and topologically secure. Due to its stego-security, this scheme is able to face attacks that take place into the "watermark only attack" framework. Its topological security reinforce its capability to face threats in other frameworks as "known message attack" or "known original attack", in the Simmons' prisoner problem. In this research work, the study of topological properties of CIS_2 is enlarged by describing this scheme as iterations over the real line, and investigating other security properties of topological nature as the Lyapunov exponent, that have been reported as important in the field of information hiding security. Results show that this scheme is able to withdraw a malicious attacker in the "estimated original attack" context too.

1 INTRODUCTION

The first fundamental work in information hiding security was realized by Cachin in the early '00s, in the context of steganography (Cachin, 2004): attempts of an attacker to make the distinction between an innocent image and a stego-content was rewritten in this article as a hypothesis testing problem. The basic properties of a stegosystem are defined by Cachin using the notions of entropy, mutual information, and relative entropy. At the same time, Mittelholzer has proposed the first theoretical framework for analyzing the security in the second category of algorithms studied by the information hiding community, namely the digital watermarking (Mittelholzer, 1999). These efforts to bring a theoretical framework for security in steganography and watermarking have been followed up by Kalker, who tries to clarify the concepts (robustness *vs.* security), and the classifications of watermarking attacks (Kalker, 2001). This work has been deepened by Furon *et al.*, who have translated Kerckhoffs' principle (Alice and Bob shall only rely on some previously shared secret for privacy), from cryptography to data hiding (Furon, 2002). They used Diffie and Hellman methodology, and Shannon's cryptographic framework (Shannon, 1949), to classify the watermarking attacks into categories, according to the type of information Eve has access to (Perez-Freire et al., 2006), namely: Watermarked Only Attack (WOA), Known Message Attack (KMA), Known Original Attack (KOA), Constant-Message Attack (CMA), and Estimated Original Attacks (EOA).

Levels of security have been recently defined in these setups. The highest level of security in WOA is called stego-security (Cayre et al., 2008), whereas topological security tends to improve the ability to withstand attacks in KMA, KOA, and CMA setups (Guyeux et al., 2010). It has been previously established that, in order to enlarge the knowledge of the level of security of a steganographic scheme, the quantity of disorder generated by the chaos of its topological security can be measured evaluating the well-known Lyapunov exponent (Martínez-Ñonthe et al., 2011; Mao and Chen, 2011; Bahi et al., 2012). The evaluation of this exponent allow to characterize the ability of the scheme to face an attacker in the context of an EOA.

The first contribution of this article consists in a relation established between how fine is a topology and the chaotic behavior of a dynamical system described with this topology. The second contribution is

the security study of a previously released stego and topologically secure steganographic scheme called $\mathcal{CIS}_2$, on a new topological space, namely the real line numbers $\mathbb{R}$. On this new space, the topological security of $\mathcal{CIS}_2$ is firstly evaluated, and its Lyapunov exponent is then computed, in order to quantify its level of disorder. Incidentally, this computation allows to measure the resistance of the $\mathcal{CIS}_2$ scheme against a category of attacks called Estimated Original Attack. This study follows a same canvas than a previous work dealing with digital watermarking, but it is conducted here for a steganographic scheme. To achieve this work, a new semi-conjugacy model must be written for the scheme $\mathcal{CIS}_2$, which is then established and proven here.

This document is organized as follows. Notions and firsts results concerning the mathematical theory of chaos are introduced in the next section. Then, in Section 3, security notions and classes of attacks under consideration in the information hiding community are recalled. In Section 4, the steganographic scheme studied in this document is presented, and the formalization allowing its topological security evaluation is given in the next section. This model is then used in Section 6 to design a new semi-conjugacy allowing its security study on a new space (Sect. 7), and its Lyapunov exponent is finally evaluated in Section 8. This paper ends by a conclusion section where our contribution is summarized and intended future researches are given.

2 THE MATHEMATICAL THEORY OF CHAOS

2.1 Notations and Terminologies

In what follows, $\mathbb{B}$ denotes the Boolean set $\{0,1\}$, S^n stands for the n^{th} term of a sequence S, V_i is for the i^{th} component of a vector V, and $[\![0;N]\!]$ is the integer interval $\{0,1,\ldots,N\}$. Furthermore, the following definitions will be used in this document.

Definition 1 The *discrete Boolean metric* is the application $\delta : \mathbb{B} \longrightarrow \mathbb{B}$ defined by $\delta(x,y) = 0 \Leftrightarrow x = y$.

Definition 2 The *vectorial negation* is the function $f_0 : \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$ defined by $f_0\left((b_0,\cdots,b_{\mathsf{N}-1})\right) = (\overline{b_0},\cdots,\overline{b_{\mathsf{N}-1}})$, where $\overline{x}$ is the negation of the Boolean x.

Definition 3 A *strategy adapter* is a sequence which elements belong into $[\![1,\mathsf{k}]\!]$, where $\mathsf{k} \in \mathbb{N}^*$. The set of all strategies having terms in $[\![1,\mathsf{k}]\!]$ is denoted by $\mathbb{S}_{\mathsf{k}}$.

Definition 4 For $k \in \mathbb{N}^*$, the *initial function* is the map $i_k : \mathbb{S}_k \longrightarrow [\![1,\mathsf{k}]\!]$ defined by $i_k\left((S^n)_{n\in\mathbb{N}}\right) = S^0$.

Definition 5 Let $k \in \mathbb{N}^*$. The *shift function* is the map $\sigma_k : \mathbb{S}_k \longrightarrow \mathbb{S}_k$ defined by $\sigma_k\left((S^n)_{n\in\mathbb{N}}\right) = (S^{n+1})_{n\in\mathbb{N}}$.

It is now possible to give some recalls in the field of the mathematical topology (Schwartz, 1980) to make this document self contained.

2.2 The Chaotic Dynamical Systems

Let $(\mathcal{X},\tau)$ be a topological space and f a continuous function on $(\mathcal{X},\tau)$.

Definition 6 f is said to be *topologically transitive* if, for any pair of open sets $U,V \subset \mathcal{X}$, there exists $k > 0$ such that $f^k(U) \cap V \neq \varnothing$.

Definition 7 $(\mathcal{X},f)$ is *regular* if the set of periodic points is dense in $\mathcal{X}$.

It is now possible to introduce the well-established mathematical definition of chaos (Devaney, 1989).

Definition 8 A function $f : \mathcal{X} \longrightarrow \mathcal{X}$ is said to be *chaotic* on $\mathcal{X}$ if it is regular and topologically transitive.

If the topological space is indeed a metric space $(\mathcal{X},d)$, then the sensibility of the system under iterations, regarding its initial conditions, can be quantified as follows.

Definition 9 f has *sensitive dependence on initial conditions* if there exists $\delta > 0$ such that, for any $x \in \mathcal{X}$ and any neighborhood V of x, there exist $y \in V$ and $n \geqslant 0$ such that $d\left(f^n(x), f^n(y)\right) > \delta$. δ is called the *constant of sensitiveness* of f.

This property is implied by both the regularity and transitivity presented above (Banks et al., 1992). And so, when f is chaotic, fundamentally different behaviors are possible for the system, and they occur in an unpredictable way.

Let us state now some basic results that surprisingly cannot be found in the literature. To simplify the presentation, some notations must be firstly introduced: $\mathcal{X}_\tau$ will stand for the topological space $(\mathcal{X},\tau)$, whereas $\mathcal{V}_\tau(x)$ is the set of neighborhoods of x for the topology τ (in unambiguous cases, we will simply use $\mathcal{V}(x)$).

Theorem 1 *Let $\mathcal{X}$ be a set, and τ,τ' two topologies on $\mathcal{X}$ such that τ' is finer than τ. Let $f : \mathcal{X} \to \mathcal{X}$, continue for both τ and τ'.*

If $(\mathcal{X}_{\tau'},f)$ is chaotic in the sense of Devaney, then $(\mathcal{X}_\tau,f)$ is also chaotic.

Proof 1 *Let ω_1, ω_2 two open sets of τ. Then $\omega_1, \omega_2 \in \tau'$, because τ' is finer than τ. As f is $\tau'-$transitive, then $\exists n \in \mathbb{N}, \omega_1 \cap f^{(n)}(\omega_2) = \varnothing$. As a consequence, f is $\tau-$transitive.*

Let us now establish the regularity of $(\mathcal{X}_\tau, f)$, i.e.*, for all $x \in \mathcal{X}$ and all $\tau-$neighborhood V of x, there exists a periodic point for f in V. Let $x \in \mathcal{X}$, and $V \in \mathcal{V}_\tau(x)$ a $\tau-$neighborhood of x. By definition of a neighborhood, $\exists \omega \in \tau, x \in \omega \subset V$. However $\tau \subset \tau'$, so $\omega \in \tau'$, and then $V \in \mathcal{V}_{\tau'}(x)$. But $(\mathcal{X}_{\tau'}, f)$ is regular, So there exists a periodic point for f in V, and the regularity of $(\mathcal{X}_\tau, f)$ is proven.*

Let us finally recall another topological quantitative property of chaos:

Definition 10 A function f is said to be *expansive* if $\exists \varepsilon > 0, \forall x \neq y, \exists n \in \mathbb{N}, d(f^n(x), f^n(y)) \geqslant \varepsilon$.

Sometimes, instead of trying to prove properties directly on the system itself, it is preferable to reduce the initial problem to another one whose characteristics are known or appear more accessible. Such a reduction tool is called, in the mathematical theory of chaos, the semi-conjugacy.

2.3 The topological semi-conjugacy

Definition 1 *The discrete dynamical system $(\mathcal{X}, f)$ is* topologically semi-conjugate *to the system $(\mathcal{Y}, g)$ if it exists a function $\varphi : \mathcal{X} \longrightarrow \mathcal{Y}$, both continuous and onto, such that:*

$$\varphi \circ f = g \circ \varphi,$$

that is, which makes commutative the following diagram (Formenti, 1998).

$$\begin{array}{ccc} \mathcal{X} & \xrightarrow{f} & \mathcal{X} \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ \mathcal{Y} & \xrightarrow[g]{} & \mathcal{Y} \end{array}$$

In this case, the system $(\mathcal{Y}, g)$ is called a factor *of the system $(\mathcal{X}, f)$.*

Various dynamical behaviors are inherited by systems factors (Formenti, 1998). They are summarized in the following proposition:

Proposition 1 *Let $(\mathcal{Y}, g)$ a factor of the system $(\mathcal{X}, f)$. Then:*

1. *for all $j \leqslant k$, $p \in Per_k(f) \Longrightarrow \varphi(p) \in Per_j(g)$, where $Per_n(h)$ stands for the set of points of period n for the iteration function h.*
2. *$(\mathcal{X}, f)$ regular $\Longrightarrow$ $(\mathcal{Y}, g)$ regular,*
3. *$(\mathcal{X}, f)$ transitive $\Longrightarrow$ $(\mathcal{Y}, g)$ transitive.*

So if $(\mathcal{X}, f)$ is chaotic as defined by Devaney, then $(\mathcal{Y}, g)$ is chaotic too.

2.4 The Lyapunov Exponent

Some dynamical systems are very sensitive to small changes in their initial conditions, which is illustrated by both the constants of sensitiveness to initial conditions and of expansiveness introduced respectively in Definitions 9 and 10. However, these variations can quickly take enormous proportions, grow exponentially, and none of these constants can illustrate that. Alexander Lyapunov has examined this phenomenon and introduced an exponent that measures the rate at which these small variations can grow.

Definition 11 Given $f : \mathbb{R} \longrightarrow \mathbb{R}$, the *Lyapunov exponent* of the system composed by $x^0 \in \mathbb{R}$ and $x^{n+1} = f(x^n)$ is defined by:

$$\lambda(x_0) = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} \ln \left| f' \left(x^{i-1} \right) \right|.$$

Consider a dynamical system with an infinitesimal error on the initial condition x_0. When the Lyapunov exponent is positive, this error will increase exponentially (situation of chaos), whereas it will decrease if $\lambda(x_0) \leqslant 0$.

Let us now recall the information hiding security framework developed this last decade.

3 INFORMATION HIDING SECURITY

In the prisoner problem of Simmons (Simmons, 1984), Alice and Bob are in jail, and they want to devise an escape plan by exchanging hidden messages in innocent-looking cover contents. These messages are to be conveyed to one another by a common warden, Eve, who over-drops all contents and can choose to interrupt the communication if they appear to be stego-contents.

In the steganography framework, in the context of the Simmons' prisoner problem, attacks have been classified in (Cayre et al., 2008) as follows:

- A *Watermark-Only Attack* WOA occurs when an attacker has only access to several watermarked contents.
- A *Known-Message Attack* (KMA) occurs when an attacker has access to several pairs of watermarked contents and corresponding hidden messages.
- A *Known-Original Attack* KOA is when an attacker has access to several pairs of watermarked contents and their corresponding original versions.

- A *Constant-Message Attack* CMA occurs when the attacker observes several watermarked contents and only knows that the unknown hidden message is the same in all contents.
- Finally, an *Estimated Original Attacks* (EOA) occurs when the attacker has access to an estimation of the original host signal, with possibly some estimation errors.

In the framework of WOA, the stego-security (Cayre et al., 2008) is relevant to evaluate the security of information hiding processes. It is the highest security level in WOA setup. To recall it, the following notations must firstly be introduced: $\mathbb{K}$ is the set of embedding keys, $p(X)$ is the probabilistic model of N_0 initial host contents, $p(Y|K_1)$ is the probabilistic model of N_0 watermarked contents. Furthermore, it is supposed in this context that each host content has been watermarked with the same secret key K_1 and the same embedding function e.

It is now possible to define the notion of stego-security:

Definition 12 (Stego-Security) In the Watermark-Only Attack framework, the embedding function e is *stego-secure* if and only if:

$$\forall K_1 \in \mathbb{K}, p(Y|K_1) = p(X).$$

In the other frameworks (KOA, KMA, and CMA), the topological security should be investigated (Friot et al., 2011). In this article, we focus more specifically on this topological security, which is recalled below.

To check whether an information hiding scheme S is topologically secure or not, S must be written as an iterate process $x^{n+1} = f(x^n)$ on a metric space $(\mathcal{X}, d)$. This formulation is always possible (Bahi and Guyeux, 2010). So,

Definition 13 (Topological Security) An information hiding scheme S is said to be topologically secure on $(\mathcal{X}, d)$ if its iterative process has a chaotic behavior according to Devaney.

Thus a data hiding scheme is secure if it is unpredictable. Its iterative process must satisfy the Devaney's chaos property and its level of topological security increases with the number of chaotic properties satisfied by it.

This new concept of security for data hiding schemes has been proposed in (Bahi and Guyeux, 2010) as a complementary approach to the existing framework. It contributes to the reinforcement of confidence into existing secure data hiding schemes. Additionally, the study of security in KMA, KOA, and CMA setups is realizable in this context. Finally, this framework can replace stego-security in situations that are not encompassed by it. In particular, this framework is more relevant to give evaluation of data hiding schemes claimed as chaotic.

In the EOA framework, the evaluation of the Lyapunov exponent, which is the subject of this research work, is relevant to quantify the level of security of steganographic processes proven to be topologically secure. Indeed, the Lyapunov exponent participates to the measurement of this topological security. In an EOA setup, the attacker has only access to estimations of the original content. With just this knowledge, he or she should not be in measure to recover any information about the secret message or the secret key. The topological security, with the two other notions of sensibility and expansiveness introduced in Definitions 9 and 10, are relevant to face attacks in this context. However, these two topological properties give no precise quantification of the security of the scheme, which justifies the consideration of the Lyapunov exponent.

4 THE STEGANOGRAPHIC SCHEME $\mathcal{CIS}_2$

To explain how to use chaotic iterations for information hiding, we must firstly define the significance of a given coefficient, and the notion of most and least significant coefficients (MSCs and LSCs).

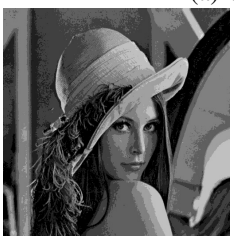
We first notice that terms of the original content x that may be replaced by terms taken from the secret message y are less important than other ones: they could be changed without be perceived as such. For a given host content x, MSCs are then ranks of x that describe the relevant part of the image, whereas LSCs translate its less significant parts. These two definitions are illustrated on Figure 1, where the LSCs correspond to the last three bits of each pixel.

The steganographic scheme $\mathcal{CIS}_2$ that generalizes the watermarking scheme based on chaotic iterations can now be recalled. In this part the following notations will be used: $x^0 \in \mathbb{B}^{\mathsf{N}}$ is the N LSCs of a given cover media C, $m^0 \in \mathbb{B}^{\mathsf{P}}$ is the secret message to embed into x^0, $S_p \in \mathbb{S}_{\mathsf{N}}$ is the *place strategy*, $S_c \in \mathbb{S}_{\mathsf{P}}$ is the *choice strategy*, and lastly $S_m \in \mathbb{S}_{\mathsf{P}}$ is the *mixing strategy*.

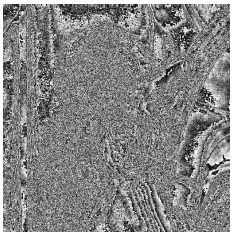
The steganographic scheme is defined by $\forall (n,i,j) \in \mathbb{N}^* \times [\![0;\mathsf{N}-1]\!] \times [\![0;\mathsf{P}-1]\!]$:

$$x_i^n = \begin{cases} x_i^{n-1} & \text{if } S_p^n \neq i \\ m_{S_c^n} & \text{if } S_p^n = i, \end{cases}$$

(a) Original.



(b) MSCs.



(c) LSCs ($\times 17$).

Figure 1: Most and least significant coefficients of Lena.

and

$$m_j^n = \begin{cases} m_j^{n-1} & \text{if } S_m^n \neq j \\ \overline{m_j^{n-1}} & \text{if } S_m^n = j. \end{cases}$$

where $\overline{m_j^{n-1}}$ is the Boolean negation of m_j^{n-1}.

The new LSCs of the stego-content are the Boolean vector $y = x^P \in \mathbb{B}^{\mathsf{N}}$.

5 TOPOLOGICAL MODEL FOR $\mathcal{CIS}_2$ AND SECURITY ANALYSIS ON $\mathcal{X}_2$

In this section is recalled the topology used in order to model the steganographic scheme $\mathcal{CIS}_2$ by a discrete dynamical system in a topological space (Devaney, 1989).

Let

$$\begin{aligned} \mathcal{F}: \quad & [\![0;\mathsf{N}-1]\!] \times \mathbb{B}^{\mathsf{N}} \times [\![0;\mathsf{P}-1]\!] \times \mathbb{B}^{\mathsf{P}} \longrightarrow \mathbb{B}^{\mathsf{N}} \\ & (k,x,\lambda,m) \longmapsto \left(\delta(k,j).x_j + \overline{\delta(k,j)}.m_\lambda\right)_{j\in[\![0;\mathsf{N}-1]\!]} \end{aligned}$$

where + and . are the Boolean addition and product operations.

Consider the phase space $\mathcal{X}_2$ defined as follow: $\mathcal{X}_2 = \mathbb{S}_N \times \mathbb{B}^{\mathsf{N}} \times \mathbb{S}_P \times \mathbb{B}^{\mathsf{P}} \times \mathbb{S}_P$, where $\mathbb{S}_N$ and $\mathbb{S}_P$ are the sets introduced in Section 2.1.

The map $\mathcal{G}_{f_0} : \mathcal{X}_2 \longrightarrow \mathcal{X}_2$ is defined by:

$$\mathcal{G}_{f_0}\left(S_p,x,S_c,m,S_m\right) =$$

$$\left(\sigma_N(S_p), \mathcal{F}\left(i_N(S_p),x,i_P(S_c),m\right), \sigma_P(S_c), G_{f_0}(m,S_m), \sigma_P(S_m)\right).$$

Then $\mathcal{CIS}_2$ can be described by the iterations of the following discrete dynamical system:

$$X^0 \in \mathcal{X}_2 \text{ and } X^{k+1} = \mathcal{G}_{f_0}(X^k).$$

By comparing $\mathcal{X}_2$ and $\mathcal{X}_1$, it has been proven in (Friot et al., 2011) that:

Proposition 1 *$\mathcal{X}_2$ has, at least, the cardinality of the continuum.*

A new distance has been defined on $\mathcal{X}_2$ as follow: $\forall X, \check{X} \in \mathcal{X}_2$, if $X = (S_p,x,S_c,m,S_m)$ and $\check{X} = (\check{S_p},\check{x},\check{S_c},\check{m},\check{S_m})$, then:

$$d_2(X,\check{X}) =$$

$$d_{\mathbb{B}^{\mathsf{N}}}(x,\check{x}) + d_{\mathbb{B}^{\mathsf{P}}}(m,\check{m}) +$$

$$d_{\mathbb{S}_N}(S_p,\check{S_p}) + d_{\mathbb{S}_P}(S_c,\check{S_c}) + d_{\mathbb{S}_P}(S_m,\check{S_m}), \text{ where:}$$

- $d_{\mathbb{B}^{\mathsf{N}}}(E,\check{E}) = \sum_{k=0}^{\mathsf{N}-1} \delta(E_k,\check{E}_k) \in [\![0;\mathsf{N}]\!]$,
- and $d_{\mathbb{S}_{\mathsf{N}}}(S,\check{S}) = \frac{9}{\mathsf{N}} \sum_{k=1}^{\infty} \frac{|S^k - \check{S}^k|}{10^k} \in [0;1]$

are respectively distances on $\mathbb{B}^{\mathsf{N}}$ and $\mathbb{S}_{\mathsf{N}}$ ($\forall \mathsf{N} \in \mathbb{N}^*$).

To demonstrate that $\mathcal{CIS}_2$ is another example of topological chaos in the sense of Devaney, it has been firstly established in (Friot et al., 2011) that,

Proposition 2 *$\mathcal{G}_{f_0}$ is continuous on $(\mathcal{X}_2,d_2)$.*

Then it has been proven that $(\mathcal{X}_2,\mathcal{G}_{f_0})$ is topologically transitive, regular, and has sensitive dependence on initial conditions. Then we have the result:

Theorem 1 *$\mathcal{G}_{f_0}$ is a chaotic map on $(\mathcal{X}_2,d_2)$ in the sense of Devaney, and consequently the scheme $\mathcal{CIS}_2$ is topologically secure.*

Another theorem about the security of $\mathcal{CIS}_2$ has been established in (Friot et al., 2011).

Theorem 2 *$\mathcal{CIS}_2$ is stego-secure.*

6 A TOPOLOGICAL SEMI-CONJUGACY BETWEEN $\mathcal{X}_2$ AND $\mathbb{R}$

In this section, by using a topological semi-conjugacy, we show that $\mathcal{CIS}_2$ modeled by $\mathcal{G}_{f_0}$ on $\mathcal{X}$ can be described as iterations on a real interval. As our researches are inspired by the work of (Bahi et al., 2012), the proofs detailed in this document will follow a same canvas. To do so, some notations and terminologies must be introduced another time.

Let $\mathcal{X}_{(\mathsf{N};\mathsf{P})} = \mathbb{S}_N \times \mathbb{B}^{\mathsf{N}} \times \mathbb{S}_P \times \mathbb{B}^{\mathsf{P}} \times \mathbb{S}_P$. In what follows and for easy understanding, we will assume

that $\mathsf{N}=3$ and $\mathsf{P}=2$. So $\mathsf{N}+\mathsf{P}=5$ and $\mathsf{NP}^2=12$. However, an equivalent formulation of the following can be easily obtained by replacing the bases 5 and 12 by any base $(\mathsf{N}+\mathsf{P})$ and (NP^2). N has only to be greater than P.

Definition 14 The function $\psi : [\![1,\mathsf{N}]\!] \times [\![1,\mathsf{P}]\!] \times [\![1,\mathsf{P}]\!] \to [\![0,\mathsf{NP}^2-1]\!]$ is defined by: $\psi\left(S_p^i,S_c^i,S_m^i\right) = (S_p^i-1)\mathsf{P}^2+(S_c^i-1)\mathsf{P}+(S_m^i-1)$.

This function aims to convert a strategy of triplets in a simple strategy of integers expressed in a different base. Obviously, ψ is a bijective function, the reverse operation will be denoted by ψ^{-1}. The three projections of ψ^{-1} are denoted by: $\psi_1^{-1}\left(\psi\left(S_p^i,S_c^i,S_m^i\right)\right) = S_p^i$, $\psi_2^{-1}\left(\psi\left(S_p^i,S_c^i,S_m^i\right)\right) = S_c^i$, and $\psi_3^{-1}\left(\psi\left(S_p^i,S_c^i,S_m^i\right)\right) = S_m^i$.

Base $\mathsf{N}=3$	Base $\mathsf{P}=2$	Base $\mathsf{P}=2$	Base $\mathsf{NP}^2=12$
S_p^i	S_c^i	S_m^i	$\psi\left(S_p^i,S_c^i,S_m^i\right)$
1	1	1	0
1	1	2	1
1	2	1	2
1	2	2	3
2	1	1	4
2	1	2	5
2	2	1	6
2	2	2	7
3	1	1	8
3	1	2	9
3	2	1	10
3	2	2	11

Table 1: Illustration of the function ψ (see Definition 14).

Definition 15 Let us define $\varphi : \mathcal{X}_{(3;2)} = \mathbb{S}_3 \times \mathbb{B}^3 \times \mathbb{S}_2 \times \mathbb{B}^2 \times \mathbb{S}_2 \longrightarrow \left[0,2^5\right[$, as follows. If $\left(S_p,E,S_c,M,S_m\right) = \left((S_p^0,S_p^1,\ldots);(E_0,E_1,E_2,E_3);(S_c^0,S_c^1,\ldots);(M_0,M_1);(S_m^0,S_m^1,\ldots)\right)$, then $\varphi\left(S_p,E,S_c,M,S_m\right)$ is the real number:

- whose integral part e is $\sum_{k=0}^{2} 2^{4-k}E_k + \sum_{k=3}^{4} 2^{4-k}M_{k-3}$, that is, the binary digits of e are $E_0\ E_1\ E_2\ M_0\ M_1$.
- whose decimal part s is equal to: $s = 0,\psi\left(S_p^0,S_c^0,S_m^0\right)\ \psi\left(S_p^1,S_c^1,S_m^1\right)\ \psi\left(S_p^2,S_c^2,S_m^2\right)\ \ldots = \sum_{k=1}^{+\infty} 12^{-k}S^{k-1}$. s is thus expressed in base 12.

φ realizes the association between a point of $\mathcal{X}_{(3;2)}$ and a real number into $\left[0,2^5\right[$. We must now translate the steganographic process $\mathcal{CIS}_2$, which is represented by $\mathcal{G}_{f_0}$ iterations on this real interval. To do so, two intermediate functions over $\left[0,2^5\right[$ denoted by e and s must be introduced.

Definition 16 Let $x \in \left[0,2^5\right[$ and:

- $e_0,\ldots,e_4$ the binary digits of the integral part of x: $\lfloor x \rfloor = \sum_{k=0}^{4} 2^{4-k}e_k$.
- $(s^k)_{k\in\mathbb{N}}$ the digits of x, expressed in base 12, where the chosen decimal decomposition of x is the one that does not have an infinite number of 11: $x = \lfloor x \rfloor + \sum_{k=0}^{+\infty} s^k 12^{-k-1}$.

e and s are thus defined as follows:

$$e: \begin{array}{ccc} \left[0,2^5\right[& \longrightarrow & \mathbb{B}^3\times\mathbb{B}^2 \\ x & \longmapsto & ((e_0,e_1,e_2);(e_3,e_4)) \end{array}$$

and

$$s: \begin{array}{ccc} \left[0,2^5\right[& \longrightarrow & [\![0,11]\!]^{\mathbb{N}} \\ x & \longmapsto & (s^k)_{k\in\mathbb{N}} \end{array}$$

We are now able to define the function g, whose goal is to translate the steganographic process $\mathcal{CIS}_2$ represented by $\mathcal{G}_{f_0}$ on an interval of $\mathbb{R}$.

Definition 17 $g : \left[0,2^5\right[\longrightarrow \left[0,2^5\right[$ is such that $g(x)$ is the real number of $\left[0,2^5\right[$ defined below:

- its integral part has a binary decomposition equal to $e'_0,\ldots,e'_4$, with $\forall i \in [\![0,2]\!]$:

$$e'_i = \begin{cases} e(x)_i & \text{if } i \neq \psi_1^{-1}\left(s^0\right) \\ e(x)_{2+\psi_2^{-1}(s^0)} & \text{if } i = \psi_1^{-1}\left(s^0\right) \end{cases}$$

  and $\forall i \in [\![3,4]\!]$:

$$e'_i = \begin{cases} e(x)_i & \text{if } i \neq \psi_3^{-1}\left(s^0\right) \\ e(x)_i + 1 \pmod 2 & \text{if } i = \psi_3^{-1}\left(s^0\right), \end{cases}$$

- whose decimal part is $s(x)^1, s(x)^2, \ldots$

In other words, if $x = \sum_{k=0}^{4} 2^{4-k}e_k + \sum_{k=0}^{+\infty} s^k\ 12^{-k-1}$, then:

$$g(x) = \sum_{k=0}^{2} 2^{4-k}\left[e_k\left(\delta(k,\psi_1^{-1}(s^0)) + 1 \pmod 2\right)\right.$$

$$+e_{2+\psi_2^{-1}(s^0)}\left(\delta(k,\psi_1^{-1}(s^0))\right)\Big]$$

$$+\sum_{k=3}^{4}2^{4-k}(e_k+\delta(k,\psi_3^{-1}(s^0)\ (\mathrm{mod}\ 2))$$

$$+\sum_{k=0}^{+\infty}s^{k+1}12^{-k-1},$$

where δ is the discrete Boolean metric introduced in Definition 1.

Numerous metrics can be defined on the set $\left[0,2^5\right[$, the most usual one being the Euclidian distance $\Delta(x,y)=|y-x|^2$. This Euclidian distance does not reproduce exactly the notion of proximity induced by our first distance d_2 on $\mathcal{X}_2$. Indeed d_2 is richer than Δ. This is the reason why we have to introduce the following metric.

Definition 18 Given $x,y\in\left[0,2^5\right[$, D denotes the function from $\left[0,2^5\right[^2$ to $\mathbb{R}^+$ defined by: $D(x,y)=D_e\left(e(x),e(y)\right)+D_s\left(s(x),s(y)\right)$, where:

$$D_e(e,\check{e})=\sum_{k=0}^{4}\delta(e_k,\check{e}_k),\ \text{ and }\ D_s(s,\check{s})=\sum_{k=1}^{\infty}\frac{|s^k-\check{s}^k|}{12^k}.$$

Proposition 3 *D is a distance on $\left[0,2^5\right[$.*

PROOF The three axioms defining a distance must be checked.

- $D\geqslant 0$, because everything is positive in its definition. If $D(x,y)=0$, then $D_e(x,y)=0$, so the integral parts of x and y are equal (they have the same binary decomposition). Additionally, $D_s(x,y)=0$, then $\forall k\in\mathbb{N}^*, s(x)^k=s(y)^k$. In other words, x and y have the same $k-$th decimal digit, $\forall k\in\mathbb{N}^*$. And so $x=y$.
- Obviously, $\forall x,y, D(x,y)=D(y,x)$.
- Finally, the triangle inequality is obtained due to the fact that both δ and $|x-y|$ satisfy it.

The convergence of sequences according to D is not the same than the usual convergence related to the Euclidian metric. For instance, if $x^n\to x$ according to D, then necessarily the integral part of each x^n is equal to the integral part of x (at least after a given threshold), and the decimal part of x^n corresponds to the one of x "as far as required". D is richer and more refined than the Euclidian distance, and thus is more precise.

φ has been constructed in order to be continuous and onto, so we obtained the following theorem:

Theorem 3 *The steganographic process $\mathcal{CIS}_2$ represented by $\left(\mathcal{G}_{f_0},\mathcal{X}_2\right)$ can be considered as simple iterations on $\mathbb{R}$, which is illustrated by the semi-conjugacy given below:*

$$\begin{array}{ccc}\left(\mathcal{X}_{(3;2)},d_2\right) & \xrightarrow{\mathcal{G}_{f_0}} & \left(\mathcal{X}_{(3;2)},d_2\right)\\ \varphi\downarrow & & \downarrow\varphi\\ \left(\left[0,2^5\right[,D\right) & \xrightarrow[g]{} & \left(\left[0,2^5\right[,D\right)\end{array}$$

In other words, $\mathcal{X}_2$ is somewhat approximately equal to $\left[0,2^{\mathsf{N}+\mathsf{P}}\right[$.

It can be remarked that the function g is a piecewise linear function: it is linear on each interval having the form $\left[\frac{n}{12},\frac{n+1}{12}\right[$, $n\in[\![0;2^5\times 12[\![$, and its slope is equal to 12. Let us justify these assessments:

Proposition 4 *The process $\mathcal{CIS}_2$ represented by g defined on $\mathbb{R}$ has derivatives of all orders on $\left[0,2^5\right[$, except on the 385 points in I defined by:*

$$I=\left\{\frac{n}{12}\ \Big/\ n\in[\![0;2^5\times 12]\!]\right\}.$$

Furthermore, on each interval of the form $\left[\frac{n}{12},\frac{n+1}{12}\right[$, with $n\in[\![0;2^5\times 12[\![$, g is a linear function having a slope equal to 12: $\forall x\notin I, g'(x)=12$.

PROOF Let $I_n=\left[\frac{n}{12},\frac{n+1}{12}\right[$, with $n\in[\![0;2^5\times 12]\!]$. All the points of I_n have the same integral part e and the same decimal part s^0: on the set I_n, functions $e(x)$ and $x\mapsto s(x)^0$ of Definition 16 only depend on n. So all the images $g(x)$ of these points x:

- Have the same integral part, which is e, except probably the bit number s^0. In other words, this integer has approximately the same binary decomposition than e, the sole exception being the digit s^0 (this number is then either $e+2^{12-s^0}$ or $e-2^{12-s^0}$, depending on the parity of s^0, *i.e.*, it is equal to $e+(-1)^{s^0}\times 2^{12-s^0}$).
- A shift to the left has been applied to the decimal part y, losing by doing so the common first digit s^0. In other words, y has been mapped into $12\times y-s^0$.

To sum up, the action of g on the points of I is as follows: first, make a multiplication by 12, and second, add the same constant to each term, which is $\frac{1}{12}\left(e+(-1)^{s^0}\times 2^{12-s^0}\right)-s^0$.

We are now able to evaluate the Lyapunov exponent of our digital watermarking scheme based on

chaotic iterations, which is now described by the iterations on $\mathbb{R}$ of the g function introduced in Definition 17.

7 TOPOLOGICAL SECURITY OF $\mathcal{CIS}_2$ ON $\mathbb{R}$

According to Theorem 1, $\mathcal{CIS}_2$ represented by the function $\mathcal{G}_{f_0}$ on $\mathcal{X}_2$ is topologically secure, that is to say $\left(\mathcal{G}_{f_0}, \mathcal{X}_2\right)$ is chaotic in the sense of Devaney. We can deduce the same property for $\mathcal{CIS}_2$ represented by the g function on $\mathbb{R}$ for the order topology. Indeed $\left(\mathcal{G}_{f_0}, \mathcal{X}_2\right)$ and $\left(g, [0, 2^5[_D\right)$ are semi-conjugate by φ as proven in the previous section. So $\left(g, [0, 2^5[_D\right)$ is a chaotic system according to Devaney, because the semi-conjugacy preserves this character (Proposition 1 in Section 2.3). However the topology generated by D is finer than the topology generated by the Euclidean distance Δ, which is the order topology. Finally, according to Theorem 1, we can affirm that the steganographic process $\mathcal{CIS}_2$ represented by g is chaotic in the sense of Devaney for the order topology on $\mathbb{R}$.

Having these assertions in mind, we can formulate the following theorem:

Theorem 2 *The steganographic process $\mathcal{CIS}_2$ represented by g on $\mathbb{R}$ is chaotic in the sense of Devaney, when the usual topology of $\mathbb{R}$ is used (the order topology).*

This result is weaker than Theorem 1, which establish the chaotic property of $\mathcal{CIS}_2$ for a finer topology. It is as if the chaos observed using usual tools like the Euclidian distance is still preserved when considering more powerful tools (higher resulution, *i.e.*, finer topologies).

The result contained in Theorem 2 is however interesting, as it confirms that the followed approach does not lead to weaker properties. Indeed, this study has taken place in a system other than the one usually considered ($\mathcal{X}_2$ instead of $\mathbb{R}$), in order to be as closed as possible to the final computer machines. By doing so, we prevent from any loss of chaotic properties when computing the scheme written in mathematical terms. However, it might be feared that the choice of a discrete mathematics approach leads to a disorder of lower quality. In other words, we have achieved to prevent from a situation of great disorder lost during the computation into machines. However, the cost of such achievement were probably to obtain a disorder of poor quality. Theorem 2 proves exactly the contrary.

8 EVALUATION OF THE LYAPUNOV EXPONENT

Let $\mathcal{L} = \left\{ x^0 \in \left[0, 2^5\right[\ \middle/ \ \forall n \in \mathbb{N}, x^n \notin I \right\}$, where I is the set of points in the real interval where g is not differentiable (as it is explained in Proposition 4). Then,

Theorem 4 *$\forall x^0 \in \mathcal{L}$, the Lyapunov exponent of $\mathcal{CIS}_2$ having x^0 for initial condition is equal to $\lambda(x^0) = \ln(12) > 0$.*

PROOF g is piecewise linear, with a slop of 12 ($g'(x) = 12$ where the function g is differentiable). Then $\forall x \in \mathcal{L}$, $\lambda(x) = \lim_{n\to+\infty} \frac{1}{n} \sum_{i=1}^{n} \ln \left| g'\left(x^{i-1}\right)\right|$ $= \lim_{n\to+\infty} \frac{1}{n} \sum_{i=1}^{n} \ln|12| = \lim_{n\to+\infty} \frac{1}{n} n \ln|12| = \ln 12$.

Remark 1 The set of initial conditions for which this exponent is not calculable is countable. This is indeed the initial conditions such that an iteration value will be a number having the form $\frac{n}{12}$, with $n \in \mathbb{N}$. Moreover, for a system having $\mathsf{N}+\mathsf{P}$ cells (a number of LSCs equal to N and a secret message to embed of width equal to P), we will find, *mutatis mutandis*, an infinite uncountable set of initial conditions $x^0 \in \left[0; 2^{\mathsf{N}+\mathsf{P}}\right[$ such that $\lambda(x^0) = \ln(\mathsf{NP}^2)$.

So, it is possible to make the Lyapunov exponent of the scheme $\mathcal{CIS}_2$ as large as possible, depending on the number of least significant coefficients of the cover media we decide to consider, and on the width of the message to embed. As proven in (Guyeux et al., 2010), a large Lyapunov exponent makes it impossible to achieve the well-known "Estimated Original Attacks" (Cayre et al., 2008).

9 CONCLUSION AND FUTURE WORK

A new quantitative evaluation for the steganographic scheme $\mathcal{CIS}_2$ is now available: its Lyapunov exponent is equal to $\ln(\mathsf{NP}^2)$, where N is the number of least significant coefficients of the cover media and P the width of the secret message to embed. This exponent allows to quantify the amplification of the ignorance on the exact initial condition (the media without watermark) after several iterations of the steganographic process. It illustrates the disorder generated by iterations of the process, reinforcing its chaotic nature. Thanks to its topological security, this scheme is already able to face an attacker in the

context of Known-Message Attack, Known-Original Attack, and Constant-Message Attack. In addition, this result implies that it is also able to resist in the context of an Estimated Original Attacks.

Using the semi-conjugacy described here, it will be possible in a future work to compare the topological behavior of $\mathcal{CIS}_2$ on $\mathcal{X}_2$ and $\mathbb{R}$, and to explore the topological security of the steganography scheme using this new topology. Then, an analogue study of the two other topologically secure schemes cited here will be conducted in order to compare these processes, being thus able to choose the best one according to the type of applications under consideration. Finally, security in steganography context will be investigated too, and topological security will be applied in this framework.

REFERENCES

Bahi, J., Friot, N., and Guyeux, C. (2012). Lyapunov exponent evaluation of a digital watermarking scheme proven to be secure. In *IIH-MSP'2012, 8-th Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, pages ***–***, Piraeus-Athens, Greece. IEEE Computer Society. To appear.

Bahi, J. M. and Guyeux, C. (2010). A chaos-based approach for information hiding security. arXiv N^o 0034939.

Banks, J., Brooks, J., Cairns, G., and Stacey, P. (1992). On devaney's definition of chaos. *Amer. Math. Monthly*, 99:332–334.

Cachin, C. (2004). An information-theoretic model for steganography. *Information and Computation*, 192:41 – 56.

Cayre, F., Fontaine, C., and Furon, T. (2008). Kerckhoffs-based embedding security classes for woa data hiding. *IEEE Transactions on Information Forensics and Security*, 3(1):1–15.

Devaney, R. L. (1989). *An Introduction to Chaotic Dynamical Systems*. Addison-Wesley, Redwood City, CA, 2nd edition.

Formenti, E. (1998). *Automates cellulaires et chaos : de la vision topologique la vision algorithmique*. PhD thesis, École Normale Suprieure de Lyon.

Friot, N., Guyeux, C., and Bahi, J. (2011). Chaotic iterations for steganography - stego-security and chaos-security. In *SECRYPT'2011, Int. Conf. on Security and Cryptography*, pages ***–***, Sevilla, Spain. To appear.

Furon, T. (2002). Security analysis. European Project IST-1999-10987 CERTIMARK, Deliverable D.5.5.

Guyeux, C., Friot, N., and Bahi, J. (2010). Chaotic iterations versus spread-spectrum: chaos and stego security. In *IIH-MSP'10, 6-th Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, pages 208–211, Darmstadt, Germany.

Kalker, T. (2001). Considerations on watermarking security. In *Multimedia Signal Processing, 2001 IEEE Fourth Workshop on*, pages 201–206.

Mao, Y. and Chen, X. (2011). An encryption algorithm of chaos based on sine square mapping. In *Proceedings of the 2011 Fourth International Symposium on Computational Intelligence and Design - Volume 01*, ISCID '11, pages 131–134, Washington, DC, USA. IEEE Computer Society.

Martínez-Ñonthe, J. A., Díaz-Méndez, A., Cruz-Irisson, M., Palacios-Luengas, L., Del-Río-Correa, J. L., and Vázquez-Medina, R. (2011). Cryptosystem with one dimensional chaotic maps. In *Proceedings of the 4th international conference on Computational intelligence in security for information systems*, CISIS'11, pages 190–197, Berlin, Heidelberg. Springer-Verlag.

Mittelholzer, T. (1999). An information-theoretic approach to steganography and watermarking. In Pfitzmann, A., editor, *Information Hiding*, volume 1768 of *Lecture Notes in Computer Science*, pages 1–16, Dresden, Germany. Springer.

Perez-Freire, L., Pérez-gonzalez, F., and Comesañ, P. (2006). Secret dither estimation in lattice-quantization data hiding: A set-membership approach. In Delp, E. J. and Wong, P. W., editors, *Security, Steganography, and Watermarking of Multimedia Contents*, San Jose, California, USA. SPIE.

Schwartz, L. (1980). *Analyse: topologie générale et analyse fonctionnelle*. Hermann.

Shannon, C. E. (1949). Communication theory of secrecy systems. *Bell Systems Technical Journal*, 28:656–715.

Simmons, G. J. (1984). The prisoners' problem and the subliminal channel. In *Advances in Cryptology, Proc. CRYPTO'83*, pages 51–67.